

DR VERONICA GREGIS (Orcid ID : 0000-0003-1876-9849)

Article type : Original Article

BPC transcription factors and a Polycomb Group protein confine the expression of the ovule identity gene *SEEDSTICK* in *Arabidopsis*

Rosanna Petrella¹, Francesca Caselli¹, Irma Roig-Villanova^{1,2}, Valentina Vignati¹, Matteo Chiara¹, Ignacio Ezquer¹, Luca Tadini¹, Martin M. Kater¹, Veronica Gregis^{1*}

¹ Università Degli Studi di Milano, Dipartimento di Bioscienze. Via Celoria 26, 20133, Milan, Italy.

² Department of Agri-Food Engineering and Biotechnology, Barcelona School of Agricultural Engineering, UPC. Esteve Terrades 8, Building 4, 08860 Castelldefels, Spain.

Contact information: *Email of corresponding author Veronica Gregis: veronica.gregis@unimi.it

Running title: BPCs, SVP and LHP1 confine *STK* expression

Keywords: MADS-box, BPCs, homeotic genes, *STK*, *LHP1*, *PRC*, *Arabidopsis thaliana*, transcription factors

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/tpj.14673](https://doi.org/10.1111/tpj.14673)

This article is protected by copyright. All rights reserved

Summary

The BASIC PENTACYSTEINE (BPC) GAGA (C-box) binding proteins belong to a small plant transcription factor family. We previously reported that BPCs of class I bind directly to C-boxes in the *SEEDSTICK* (*STK*) promoter and the mutagenesis of these *cis*-elements affects *STK* expression in the flower. The MADS-domain factor SHORT VEGETATIVE PHASE (SVP) is another key regulator of *STK*. Direct binding of SVP to CArG-boxes in the *STK* promoter are required to repress its expression during the first stages of flower development. Here we show that BPCs of class II directly interact with SVP and that MADS-domain binding sites in the *STK* promoter region are important for the correct spatial and temporal expression of this homeotic gene. Furthermore, we show that BPCs of class I and II act redundantly to repress *STK* expression in the flower, most likely by recruiting TERMINAL FLOWER 2/LIKE HETEROCHROMATIN PROTEIN 1 (TFL2/LHP1) and mediating the establishment and the maintenance of H3K27me3 repressive marks on the DNA. We investigate the role of LHP1 in the regulation of *STK* expression. Besides providing a better understanding of the role of BPC transcription factors in the regulation of *STK* expression, our results suggest the existence of a more general regulatory complex composed of BPCs, MADS-domain factors and PRCs, that cooperate to regulate gene expression in reproductive tissues. We believe that our data along with the molecular model herein described could provide significant insights for a more comprehensive understanding of gene regulation in plants.

Introduction

Transcription factors (TFs) are regulators of gene expression; they act at multiple levels to orchestrate developmental processes. TFs bind specific DNA sequences and they can cooperate through genetic and epigenetic mechanisms. TFs act in multimeric complexes that can include members of different TFs families and other proteins. The composition of these complexes determines their binding specificity and their activity on target gene regulation (Martinez and Rao, 2012). Although in the last decades different classes of plant TFs have been characterised, the molecular mechanisms by which they act and the complexes they are part of, are yet to be fully understood. Recently, a new class of transcription factors, named BASIC PENTACYSTEINE/ BARLEY B RECOMBINANT (BPC/BBR), has been identified (Santi et al., 2003). BPCs bind the RGARAGRRA consensus site, also called GAGA or C-box, to regulate their target genes (Meister et al., 2004; Kooiker et al., 2005; Simonini et al., 2012; Simonini and Kater, 2014; Hecker et al., 2015; Mu et al., 2017; Shanks et al., 2018; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). BPCs have been described in different plant species including monocots (*Oryza sativa* (rice) and *Hordeum vulgare* (barley)) and dicots (*Glycine max* (soy-bean) and *Arabidopsis thaliana*) (Sangwan and O'Brian, 2002; Santi et al., 2003; Kooiker et al., 2005; Monfared et al., 2011; Berger and Dubreucq, 2012; Simonini et al., 2012; Simonini and Kater, 2014; Hecker et al., 2015; Mu et al., 2017; Xiao et al., 2017; Shanks et al., 2018; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). In *Arabidopsis*, BPCs are divided into three subfamilies: class I (containing *BPC1* to *BPC3*), class II (containing *BPC4* to *BPC6*), and class III (containing only *BPC7*) (Meister et al., 2004; Monfared et al., 2011). Except for *BPC5*, which is a pseudogene, all the other BPCs are ubiquitously expressed. Combinations of multiple *bpc* mutants show strong phenotypes with a wide range of defects, addressing an important role during plant development (Monfared et al., 2011).

Previously, we have identified the MADS-box gene *SEEDSTICK* (*STK*) as a direct target of BPCs belonging to the class I (Kooiker et al., 2005; Simonini et al., 2012). *STK* is specifically expressed during ovule and seed development and has a wide range of functions in these tissues (Favaro et al., 2003; Pinyopich et al., 2003; Brambilla et al., 2007; Losa et al., 2010; Mizzotti et al., 2014; Mendes et al., 2016; Balanzà et al., 2016; Ezquer et al., 2016; Herrera-Ubaldo et al., 2019).

During carpel development, *STK* expression is confined to placental tissues and ovule primordia; in mature ovules, it is expressed strongly in the funiculus and in integuments that will later form the seed coat (Mizzotti et al., 2014). *STK* acts redundantly with two other MADS-box factors named *SHATTERPROOF 1 (SHP1)* and *SHATTERPROOF 2 (SHP2)* in the determination of ovule identity (Favaro et al., 2003; Pinyopich et al., 2003). BPCs of class I form homo- and heterodimers and bind C-boxes in the promoter of *STK* inducing DNA loop formation (Kooiker et al., 2005). C-boxes are important for *STK* regulation since mutations in these sequences result in the ectopic expression of the homeotic gene in the flower (Kooiker et al., 2005; Simonini and Kater, 2014). The MADS-domain factor SHORT VEGETATIVE PHASE (SVP) is another key regulator of *STK*. SVP acts redundantly with *APELATA1 (AP1)* and *AGAMOUS-LIKE24 (AGL24)* to repress *STK* expression during early stages of flower development, by binding directly to its promoter (Simonini et al., 2012; Gregis et al., 2013). Furthermore, BPCs of class I and SVP directly interact to repress *STK* expression in the floral meristem, and C-boxes are important to facilitate the binding of SVP to the *STK* promoter region (Simonini et al., 2012).

Recently, members of the BPCs family have been shown to be implicated in the recruitment of histone-modifying complexes that can inactivate gene expression, like the Polycomb Repressive Complexes (PRCs) (Hecker et al., 2015; Mu et al., 2017; Xiao et al., 2017; Roscoe et al., 2019; Wu et al., 2019). BPCs of class II directly interact with LHP1, a component of plant PRC1 that is associated with genes marked by trimethylation of histone H3 lysine 27 (H3K27me3) (Hecker et al., 2015). Furthermore, LHP1 also acts as a component of PRC2 to establish H3K27me3 and has a role in maintaining this mark at PRC2 target genes (Zhang et al., 2007b; Derkacheva et al., 2013; Wang et al., 2016). Interestingly, it was demonstrated that SVP can form heterodimers with LHP1 to modulate H3K27me3 deposition on the *SEPALLATA3 (SEP3)* locus (Liu et al., 2009). Furthermore, BPCs can physically interact with the PRC2 subunit SWINGER (SWN) to repress the expression of their target *ABSCISIC ACID INSENSITIVE4 (ABI4)* during root development by the trimethylation of Histone H3 Lysine 27 (Mu et al., 2017); moreover a close proximity of BPC6 with *VERNALIZATION2 (VRN2)* has been reported (Hecker et al., 2015).

Here we clarify the molecular mechanisms by which BPCs of class II and SVP act in the regulation of *STK* expression. We show that MADS-domain binding sequences in the *STK* promoter region

are important for the correct spatial and temporal expression of the ovule identity gene. Our data indicate that both BPCs of class I and II redundantly control the expression of *STK*, by modulating the deposition and/or the maintenance of H3K27me3 marks. Our results provide insights into the molecular mechanisms that drive transcription regulation in plants and investigate the involvement of a protein complex in which BPCs, MADS-domain factors and LHP1 can cooperate to orchestrate the expression of homeotic genes during plant development.

Results

***STK* is deregulated during flower development in the *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutant**

To gain more insights into the role of class I and class II BPCs in the regulation of *STK* expression during flower development we generated the *bpc1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant (henceforth called *bpcV*). In contrast with the previously published quintuple mutant, this mutant includes the *bpc1-2* allele which leads to a complete knock-out of the gene (Monfared et al., 2011; Simonini and Kater, 2014). The contribution of class I and II BPCs to the correct regulation of *STK* was analysed by *in-situ* hybridisation assays (Figure 1). In wild-type plants, *STK* expression was confined to ovules and the placenta and was never observed in flowers before stage 8 neither in inflorescences nor in floral meristems (Figure 1a, b). The knock-out of all the BPCs of class I (Figure 1c, d) or class II (Figure 1e, f) did not affect *STK* expression in the flower. In contrast, in the *bpcV* mutant the expression of *STK* was not only observed in ovules and placenta, but also in floral meristems, young flowers and developing petals (Figure 1h). *STK* expression was also detectable in flower organ primordia (Figure 1g). The *in-situ* hybridisation controls, using a H4 gene- specific probe (confirming the integrity of the tissue) and a *STK* sense probe (Figure S1; Fobert et al., 1994; Favaro et al., 2003), were performed to confirm the *in-situ* data. These results clearly demonstrated the redundant role that class I and II BPCs have in the regulation of *STK* expression during flower development.

Phenotypical characterization of the *bpcV* mutant and 35S:*STK* lines

To further investigate the role of BPCs in plant development, we performed a phenotypical analysis of the *bpcV* mutant. The quintuple mutant plants were shorter when compared to wild-

type plants and were characterized by both vegetative and reproductive defects (Figure 2a, b and Figure S2b and c). The knockout of the five BPC genes caused a drastic phenotype in the siliques. In wild-type, upon successful fertilization, from 3 to 12 days after pollination (dap), the siliques elongate to reach their maximum length. In contrast, in the *bpcV* mutant no silique elongation was registered (Figure 2b). As also pointed out in Monfared et al. (2011), these results suggest that BPCs are involved in different aspects of plant development.

To analyse the phenotypic effects of the deregulation of *STK* and to compare them with the phenotypes observed in the *bpcV* mutant, we transformed wild-type plants with a chimeric gene construct in which the CDS of *STK* was fused to the Cauliflower mosaic virus (CaMV) 35S promoter (Favaro et al., 2003). *STK* expression was analysed by quantitative Real-Time PCR in three lines, where we could detect statistically significant upregulation of *STK* expression (Figure S2a). The line that showed the highest upregulation (henceforth called 35S:STK) was propagated and in the next generations used for further analysis. Intriguingly, also this plant was shorter compared to the wild-type (Figure 2a) and showed defects that phenocopy the *bpcV* mutant, including the silique phenotype (Figure 2b) which is consistent with an upregulation of *STK* in the *BPCV* mutant.

The MADS-domain factor *STK* is a master player in ovules and seeds development (Favaro et al., 2003; Pinyopich et al., 2003; Mizzotti et al., 2014; Ezquer et al., 2016).

To determine whether the constitutive expression of *STK* affected seed development, seed area was analysed in 35S:STK plants and *bpcV* mutants. As a control, wild-type, *stk* and *arf2-8* seeds were used. Our results confirmed that *stk* had smaller seeds, as previously reported by Pinyopich et al. (2003) whereas *arf2-8* seeds were larger (Schruff et al., 2006). Interestingly, both *bpcV* and 35S:STK plants showed a wider seed area when compared to the wild-type and the *stk* mutant, even though *bpcV* mutant seeds were larger than those of 35S:STK (Figure 2c). Our results support the hypothesis that BPCs regulate *STK* expression in the gynoecium and in seeds.

BPCs of class II interact with SVP

Previously, we have shown that class I BPCs act together with SVP in the control of *STK* expression (Simonini et al., 2012). The *in-situ* analysis (Figure 1) suggests that BPCs of class II have an important role in regulating *STK* expression. To understand whether BPCs of class II (BPC4 and BPC6) interact with SVP, different protein interaction assays were performed.

We confirmed by yeast two-hybrid assays and bimolecular fluorescence complementation assays (BiFC) in tobacco leaves (*Nicotiana benthamiana*) that BPC4 and BPC6 can form homo- and heterodimers (Wanke et al., 2011; Figure S3a, c). Furthermore, we showed, using yeast two-hybrid assays, that both BPC4 and BPC6 can interact with SVP (Figure 3a). To confirm the interactions between SVP and the BPC4 and BPC6 factors, a co-Immunoprecipitation (Co-IP) assay was performed, using SVP-GFP in combination with BPC4-RFP and BPC6-RFP fusion proteins, transiently co-expressed under the control of the Cauliflower mosaic virus (CaMV) promoter in *Nicotiana benthamiana* leaves. These Co-IP experiments all revealed co-precipitation of the BPC and SVP proteins, respectively (Figures 3b and Figure S3b), suggesting that BPC4 and BPC6 are able to interact with SVP *in vivo*.

Further validation of these results *in planta* was obtained by BiFC assays in tobacco (*Nicotiana benthamiana*) leaves. The combination SVP-YFP^N BPC4-YFP^C showed a clear nuclear interaction between BPC4 and SVP (Figure 3c). All the other combinations that were tested (BPC4-YFP^N SVP-YFP^C, SVP-YFP^N BPC6-YFP^C, BPC6-YFP^N SVP-YFP^C) resulted in an interaction in the cytoplasm (Figure 3c, all the control experiments are reported in Figure S3-S5). Although this result was unexpected, Immink et al. (2002) previously showed that some MADS-domain proteins need to dimerise with another MADS-domain factor for their nuclear localisation. SVP interacts with the MADS-domain protein AP1 during floral development and therefore it might facilitate the nuclear location of SVP-BPC dimers (Pelaz et al., 2002; de Folter et al., 2005). To test this hypothesis, we co-expressed SVP-BPC4 and SVP-BPC6 dimers with an AP1-RFP fusion protein in tobacco leaves. As shown in Figure 3d, the presence of AP1 facilitates the nuclear localisation of the BPC4-SVP and BPC6-SVP dimers. To determine whether BPCs of class II could directly interact with AP1, a BiFC interaction assay was performed which showed no interaction between BPCs of class II and AP1, as reported in Figure S3c. Taken together these results clearly show that AP1 is sufficient for the translocation of class II BPCs-SVP heterodimers to the nucleus.

Finally, we also analysed the subcellular localization of BPC4/6-RFP, SVP-GFP and AP1-RFP in tobacco leaves (*Nicotiana benthamiana*). As shown in Figure S4b, BPC4, BPC6 and AP1 all localized in the nuclei whereas SVP localization was registered in the cytoplasm as well, suggesting that in the BiFC assays described above, SVP is in most combinations tested the critical factor for cytoplasmatic localization.

Molecular mechanism of SVP-class I BPCs binding to the regulatory region of *STK*

To clarify the mechanism by which BPCs and SVP interact with the *STK* regulatory region, we performed a series of Chromatin Immunoprecipitation (ChIP) experiments in different mutant backgrounds.

As shown in Figure 4a, SVP binds CARG-boxes that are surrounded by C-boxes in the regulatory region of *STK*. As previously shown, SVP, AP1 and AGL24 redundantly determine the identity of the floral meristem through direct repression of floral homeotic genes (Gregis et al., 2006; Gregis et al., 2008; Gregis et al., 2009). In fact, in the *svp agl24 ap1-12* triple mutant, *STK* is ectopically expressed in floral meristems and young flowers (Simonini et al., 2012). To determine whether SVP, AP1 and AGL24 are required for BPCs binding to the promoter of *STK*, three independent ChIP assays using specific antibodies against class I BPCs were performed. The experiments were conducted using *svp agl24 ap1-12* triple mutant inflorescences. Furthermore, inflorescences from wild-type and *bpc1-2 bpc2 bpc3* triple mutant plants were used as a positive and negative control, respectively. In our ChIP experiments, no enrichment was detected in the *svp agl24 ap1-12* triple mutant in the region containing C-box 12 and the region containing C-box 4 and 5 (region B) (Figure 4b). These results demonstrate that SVP, AP1 and AGL24 are necessary for the binding of class I BPCs to the *STK* promoter.

Subsequently, the role of class I and class II BPCs in the binding of SVP to the promoter of *STK* was investigated by crossing the *bpcV* mutant, described above, with *pSVP:SVP-GFP svp* plants. In subsequent generations, plants homozygous for the *svp* and *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutations containing the *pSVP:SVP-GFP* construct were selected. ChIP experiments using commercial antibodies against GFP were performed. Inflorescences from *pSVP:SVP-GFP svp*

plants were used as a positive control, whereas wild-type was used as a negative control. An enrichment was detected when binding to the consensus regions for SVP was tested in the *bpcV* mutant background (Figure 4c). These results suggest that BPCs of class I and class II are not necessary for SVP binding to *STK* promoter.

Taken together, the results obtained by these ChIP assays are consistent with a model where SVP binds the *STK* promoter independently of BPCs, whereas BPCs of class I require MADS-domain factors for the correct binding to the *STK* regulatory region.

CARG-boxes drive the correct temporal and spatial expression of *STK* and are important for SVP and BPCs of class I binding to the promoter of *STK*

To further characterise the role of SVP in *STK* regulation, we decided to perform a functional characterisation of the CARG-boxes contained in *STK* regulatory region; these regions were identified based on the MADS-domain factor consensus binding sequences located in the *STK* locus where SVP binding was detected by ChIP-seq (Gregis et al., 2013). Considering our previous experiments using a *STK* promoter with mutated C-boxes or CARG-boxes (Simonini et al., 2012; Mendes et al., 2016), we suspected that the 12 CARG-boxes in the regulatory region of *STK* could be redundant. Therefore, a mutated version of the *STK* promoter was used in which 11 out of the 12 CARG-boxes were altered, considering the following criteria: (i) preserving the DNA conformation, introducing only 4 to 5 transitions to each consensus; (ii) avoiding the mutation of C-boxes; (iii) preventing the formation of new CARG-boxes (see Table S1). The mutagenized *STK* promoter was fused to the *uidA* reporter gene that encodes for beta-glucuronidase (GUS) and the resulting *pSTK_CARGm*:GUS construct was used to transform Arabidopsis wild-type plants and *pSVP*:SVP-GFP *svp* plants. As a positive control, the wild-type *STK* promoter (*pSTK_CARGwt*:GUS), which drives specific expression in the placenta and all stages of ovule development, was used (Figure 5a). Out of the 39 plants transformed with the *pSTK_CARGwt*:GUS construct, 36 showed a correct spatial and temporal expression of the GUS reporter, reflecting the endogenous expression of *STK* (Figure 5b-d), whereas the other three plants did not show any GUS activity. In contrast, out of 41 plants transformed with the *pSTK_CARGm*:GUS construct, two plants did not show any GUS activity, whereas 14 (36%) showed strong deregulation of GUS

expression while the remaining 64% showed a correct expression of the reporter. Interestingly, GUS expression was extended also in the inflorescence and floral meristems (Figure 5e, f) and in all the floral organs (Figure 5f). These results support the idea that the CArG-boxes in the *STK* promoter are important for the correct expression of this MADS-box gene, although the possibility that other transcription factor binding sites (cis-elements) might be involved in the regulation of *STK* cannot be excluded.

To assess whether MADS-domain binding sites on the *STK* promoter are necessary for SVP binding, we performed ChIP experiments using antibodies against GFP and inflorescences of *pSVP:SVP-GFP svp* plants with *pSTK_CArGm:GUS* that showed deregulation of the reporter. The wild-type endogenous *STK* promoter was used as a positive control, whereas inflorescences of *pSTK_CArGm:GUS* plants without SVP:GFP served as a negative control. Specific primers were designed to discriminate between the endogenous (wild-type) and the mutagenized promoter (see Experimental procedures and Table S2). These experiments showed that no enrichment was detected when binding to the mutated region was tested (Figure 5h), which suggests that CArG-boxes in the promoter of *STK* are important for the binding of SVP.

To clarify the partial penetrance of the GUS deregulation phenotype, we investigated the binding of SVP to the mutated *STK* promoter in the *pSTK_CArGm:GUS* lines that did not show deregulation of the reporter. Interestingly, in these plants binding of SVP-GFP was still observed (Figure 5i), providing clear evidence that the presence or absence of SVP binding determined the correct spatial expression profile of *STK*.

To further investigate the role of SVP and BPCs in the regulation of *STK* expression, binding of class I BPCs to the *pSTK* promoter with the mutated CArG-boxes was tested using the *pSTK_CArGm:GUS* lines that showed deregulation of the reporter and using antibodies against BPCs. As a positive control, the endogenous region of the *STK* promoter was used, whereas as negative control antibodies against HA were used. No enrichment was detected when BPCs binding to the mutated region was tested, suggesting that mutagenesis of CArG-boxes abolished BPCs of class I binding (Figure 5l). Collectively, these results indicate that CArG-boxes are important for the correct spatiotemporal regulation of *STK* expression. Moreover, these

experiments confirm that SVP binding is necessary for the recruitment of BPCs of class I to the *STK* promoter.

The expression of *STK* is influenced by epigenetic modifications

Recently, a novel role for BPCs in the regulation of target gene expression by the recruitment of PRCs, has been reported (Hecker et al., 2015; Mu et al., 2017; Xiao et al., 2017; Shanks et al., 2018; Roscoe et al., 2019; Wu et al., 2019). The presence of H3K27me3 is mainly correlated with gene silencing. However, in *Arabidopsis* H3K27me3 marks are also found to be enriched in genes with tissue-specific expression patterns, suggesting that this epigenetic mark is modulated in response to developmental cues (Zhang et al., 2007a). Interestingly, the locus of *STK* shows strong coverage of H3K27me3 deposition at least in seedlings (Turck et al., 2007; Lafos et al., 2011; Li et al., 2015).

To understand whether the *STK* locus was also decorated with H3K27me3 marks in reproductive tissues, we performed ChIP experiments using wild-type inflorescences. We included in these experiments also the analysis of *bpcV* inflorescences to investigate a possible role for BPCs in the deposition of H3K27me3 marks on the *STK* locus. Two regions were tested: region 1, located in the first *STK* intron, and region 2 immediately after the stop codon of the gene, as illustrated in Figure 6a. ChIP experiments were performed using specific antibodies against H3K27me3 and analysed by quantitative Real-Time PCR. The *AT2G22560* and *AGAMOUS* loci were used as a negative and positive control for H3K27me3 marks, respectively (Li et al., 2015). Our results showed that the *STK* locus was also decorated with H3K27me3 in inflorescence tissue. Interestingly, in *bpcV* mutant inflorescences, a reduction of H3K27me3 deposition was detected in both the two selected *STK* regions (Figure 6b). These results were consistent with the observed ectopic expression of *STK* in the *bpcV* background (Figure 1g, h) and suggest an active role of BPCs in the establishment of repressive epigenetic marks.

LHP1 is involved in *STK* regulation during flower development

It was shown that SVP and BPC6 are both able to interact with the PRC1 factor LHP1 (Liu et al., 2009; Hecker et al., 2015). Furthermore, BPC6 possesses an Alanine Zipper-Like Coiled-Coil domain at the N-terminal region, which was shown to be essential for homo- or hetero-dimerization with other members of the BPC class II family as well as for the interaction with LHP1. The Alanine Zipper-Like Coiled-Coil domain is not present in class I and class III BPCs (Berger and Dubreucq, 2012; Hecker et al., 2015).

LHP1 recognises loci marked by H3K27me3 *in vivo*, acting as part of a mechanism that represses the expression of PRC2 targets (Turck et al., 2007). Furthermore, it has been recently reported that LHP1 could directly interact with several members of the PRC2 protein family (Derkacheva et al., 2013; Wang et al., 2016) to facilitate their recruitment to target genes.

Mutation of the LHP1 locus resulted in pleiotropic effects due to the deregulation of several genes during plant development (Larsson et al., 1998). To address the role of LHP1 in the regulation of *STK* during flower development, its expression was analysed in the *lhp1* mutant background by *in-situ* hybridisation (Figure 7). In line with our hypothesis, the knock-out of *LHP1* deregulated *STK* expression in the flower, showing expression in floral and inflorescence meristems, as well as in young flowers (Figure 7c) whereas the expression of *STK* in mature flowers was not altered (Figure 7d). *STK* expression was also analysed in *lhp1* inflorescences by quantitative Real-Time PCR. The MADS-box gene *STK* is upregulated in the *lhp1* mutant background, as shown in Figure 7e.

ChIP-on-chip data presented by Turck et al. (2007) suggest that LHP1 is associated to the *STK* genomic region in seedlings. To validate the high throughput data and analyse the association to the *STK* locus in reproductive tissues we performed a ChIP assay, collecting inflorescences from *pLHP1:LHP1-GFP lhp1* plants (Kotake et al., 2003). As previously done for H3K27me3 ChIP experiments, we tested region 1 and region 2 since they have been previously reported to be associated to LHP1 in seedlings (Turck et al., 2007). We confirmed LHP1 association to the *STK* locus in Region 2, close to the 3'UTR of the homeotic gene (Figure 7f). Collectively, these results support a role of LHP1 in the modulation of *STK* expression in the flower.

Genome-wide analysis of BPCs and MADS-domain factor binding site locations

DNA affinity purification sequencing (DAP-seq) is a transcription factor (TF)-binding site discovery assay which combines next-generation sequencing of a genomic DNA library with affinity-purified TFs (Bartlett et al., 2017). In publicly available repository of *A. thaliana* transcription factor binding profiles (http://neomorph.salk.edu/dap_web/pages/index.php; O'Malley et al., 2016) patterns of genome-wide co-enrichment of MADS-domain and BPCs transcription binding sites were investigated. Average profiles for the two families were reconstructed by using a simple consensus method (see Experimental procedures). Overlap of genomic regions associated with DAP-seq peaks of MADS-domain and BPCs TFs families was used as a proxy to investigate possible interactions.

A highly significant over-representation of overlapping peaks (p -value hypergeometric $\leq 3.5e-4$) was observed which can be considered as an indication of a possible direct interaction between MADS-domain and BPC proteins. Of note, our analysis of the complete dataset of O'Malley et al. (2016), which provides DAP-seq data for more than 500 TFs, belonging to 41 distinct transcription factor families, suggests that overall only 4 additional families of transcription factors show significant levels of overlap with DAP-seq peaks of members of the BPC family (REM, C2C2gata, SRS and Trihelix, Data S1). As outlined in Figure S6, overlap with DAP-seq peaks associated with transcription factors of the MADS-domain family accounts for 19.6% of the total number of significantly overlapped peaks. All in all, we believe that these data are consistent with a model where BPCs can interact with a restricted set of TFs families, which is not limited to but is very likely to include members of the MADS-domain family. *In silico* prediction of enriched sequence motifs is largely concordant with this model (Figure 8). In fact, when *de-novo* reconstruction of enriched motifs is performed we observe: i) a strong enrichment in CARG-box like motifs in genomic regions that are bound by MADS-domain factors; ii) a strong enrichment of C-boxes like motifs in genomic regions associated with BPC DAP-seq peaks; iii) a strong enrichment of both type of motifs (CARG-boxes and C-boxes) when regions containing coincident DAP-seq peaks are considered.

Consistent with this model, analyses of a carefully selected collection of publicly available genome-wide binding profiles *in vivo*, as determined by ChIP-sequencing, showed highly

significant levels of overlap between BPCs (BPC1 and BPC6; Xiao et al., 2017; Shanks et al., 2018) and MADS-domain factors (SVP, SOC1, FLC and FLM; Mateos et al., 2015; Immink et al., 2012; Posé et al., 2013) in vegetative tissues. Values of overlap were varying between 18% and 36% for BPC6, whereas percentages were between 3% and 10% for BPC1 (Data S2 and Supplementary Figure S7).

To further investigate possible biological pathways regulated by MADS-BPCs complexes, genomic regions associated with overlapped MADS-domain and BPCs DAP-seq peaks were annotated by the means of the “AnnotatePeaks” program from the Homer suite (Heinz et al., 2010). A total of 519 candidate target genes was obtained, which were subjected to functional enrichment analyses using the DAVID program (Huang et al., 2009; Data S3).

Coincident MADS-BPCs and DAP-seq peaks were subsequently cross-referenced with H3K27me3 ChIP-seq peaks (Lafos et al., 2011), to gain a better insight on potential MADS-BPCs target genes that are co-regulated *in vivo*. A list of 93 candidate genes was obtained (Data S3). Interestingly, functional enrichment analysis of this set of genes resulted in a significant enrichment of transcription factor encoding genes (GO term “DNA-binding transcription factor activity”) for both lists: genes associated with MADS-BPCs peaks and/or with MADS-BPCs and H3K27me3. These results support the idea that MADS-BPC-PRC complexes play a pivotal role in the regulation of master players in development as shown in Data S3.

Discussion

Our knowledge on the molecular mechanisms controlling gene expression in plants is still fragmented and needs further study. Here we used the ovule identity gene *STK*, which is specifically expressed in Arabidopsis placenta, ovules and seeds, as a model system to investigate the regulation of homeotic genes expression. Previously, we showed that the MADS-domain factors SVP, AGL24 and AP1 and the class I BPC transcription factors repress *STK* expression during early stages of flower development (Kooiker et al., 2005; Simonini et al., 2012; Gregis et al., 2013). Information concerning the role of the MADS-domain factors in *STK* regulation were derived from the analysis of the *agl24 svp ap1-12* triple mutant, whereas the role of BPC factors

in *STK* regulation was mainly investigated by mutagenesis of multiple BPCs binding sites (C-boxes) in the *STK* regulatory region. Mutation of those sites caused indeed strong deregulation of *STK* during flower development, suggesting that C-boxes and therefore BPC factors are important for *STK* regulation (Simonini et al., 2012). Even though these data are very interesting, mutating cis-elements still provides indirect evidence for the role of BPCs in *STK* regulation. This study provides now deeper insights in the interplay between Class I and II BPC proteins and MADS-domain factors in the ovule specific regulation of *STK*.

The roles of BPCs and MADS-domain factors containing complexes in the regulation of STK

We investigated the role of BPC factors in the regulation of *STK* by generating the *bpc1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant (*bpcV*). In this mutant *STK* expression was deregulated and its transcripts were detected by *in-situ* hybridisation in the floral meristem and floral organs, confirming the importance and redundant role of BPCs, belonging to both class I and class II, in *STK* regulation. Intriguingly, ChIP experiments showed that in the *bpcV* mutant background SVP could still bind the *STK* promoter, suggesting that SVP can bind DNA independently of BPC factors. Previously, we reported that the mutagenesis of BPC binding sites (C-boxes) in the *STK* promoter affected SVP binding to the DNA. This discrepancy could not be due to the fact that nearby C-box mutations influence CArG-box affinity since control experiments ruled this option out (Simonini et al., 2012). However, we cannot exclude the possibility that mutagenesis of C-box elements in the promoter of *STK* introduced structural changes that altered the binding affinity of regulatory elements causing loss of binding of SVP or of unknown co-factors important for SVP binding to the *STK* promoter.

The *in silico* analyses, using the open-access database Jaspar (Khan et al., 2018), of the regulatory region of *STK* revealed binding motifs for different transcription factors of Arabidopsis. Interestingly, motifs recognized by HD-ZIP homeobox domain factors, C2H2 zinc finger Dof domain factors, Basic helix-loop-helix factors (bHLH), Beta-Hairpin-Ribbon AP2 MBD_like and GATA-type zinc fingers were detected. All these TFs families could indeed be part of the regulatory machinery for the correct spatiotemporal expression of *STK*. This scenario might also explain why the deregulation phenotype of the lines with the mutated *STK* promoter (*pSTK_CArGm:GUS*) was not fully penetrant (36% of the plants showed GUS expression outside

the normal *STK* expression domain). The binding of the *STK* repressive complex might not only be dependent on the association of SVP with the promoter but other cofactors are expected to be involved in recruiting the complex to the DNA. In the absence of SVP binding, the interaction of the complex with the mutated *STK* promoter could be less stable and influenced by for instance fluctuations in environmental conditions.

It remains important to underline that all our experiments demonstrate that BPC and MADS-domain factors are together essential for the correct expression of *STK* and that, binding of SVP alone is not *per se* sufficient to repress *STK* expression in the floral meristems.

We previously revealed that BPCs of class I can interact with each other (Simonini et al., 2012); moreover, BPCs of class II form homo and heterodimers with members of class I (Wanke et al., 2011; Simonini et al., 2012 and Figure S3). BPC protein-protein interactions studies suggest that BPC factors of class I and II can act synergistically and redundantly to regulate the expression of their targets as we demonstrated for *STK*. An example has been provided by Mu et al. (2017), who showed that mutations in BPCs of class I and II increased *ABI4* expression in roots.

To further investigate the molecular and functional relationships between the MADS-domain factor SVP and BPCs of class II, we tested their ability to form heterodimers *in planta*. We revealed that SVP interacts with BPC4 and BPC6, but the dimers are mainly retained in the cytoplasm. An interesting observation was that the MADS-domain protein AP1, an interactor of SVP, facilitated the co-localisation of SVP-BPC4 (and BPC6) to the nucleus. These data further clarify the role of AP1 in the regulation of *STK* (Simonini et al., 2012).

The Role of BPCs during seed development

Pinyopich et al. (2003) reported that *STK* has also a role during seed development since the *stk* mutant presented smaller seeds compared to the wild-type.

In this work we found that BPCs can restrict the expression of *STK* in certain spatiotemporal window since in the *bpcV* mutant the expression of *STK* was extended to other tissues in floral organs. The analysis of the *35S:STK* line here presented, further explored the effects of the deregulation of *STK* throughout reproductive development. The defects in seeds size registered in the *bpcV* and *35S:STK* suggest that BPCs might control *STK* expression later during

development, in seeds. The observation that in the *bpcV* mutant seeds are bigger than in 35S:STK lines, suggests an additive role of BPCs during seed development. It will be interesting to further investigate the role of BPCs in such an important aspect of plant development which has enormous implications in agronomical species.

BPCs of class II and SVP recruit LHP1 for the regulation of *STK*

Farkas et al. (1994) have first characterised the GAGA Associated Factor of *Drosophila melanogaster* (dGAFs). Even though GAFs and BPCs are phylogenetically unrelated, they present several similarities. BPCs can bind to (GA)_n sequences (Berger and Dubreucq, 2012) to control the expression of their targets (Meister et al., 2004; Berger et al., 2011; Simonini et al., 2012; Simonini and Kater, 2014; Mu et al., 2017; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). They also present a highly conserved zinc finger like DNA-binding domain, similar to *Trl* of *Drosophila* (Wanke et al., 2011). Interestingly, cooperative binding of BPC1 proteins to GA-rich motifs in the *STK* promoter region leads to condensation and looping of DNA (Kooiker et al., 2005), similar to what has been described for dGAF from *Drosophila*. Recent works in *Arabidopsis* revealed an intriguing interaction among BPCs and Polycomb group proteins, similar to those described in animals for dGAF, which can cooperate with Polycomb Group factors (PcG) to repress gene expression (Horard et al., 2000). PcG complexes have paramount roles in cell fate determination and differentiation both in plants and in animals. These proteins have been identified in *Drosophila* more than 40 years ago as key repressors of homeotic genes (Hox) throughout embryonic development (Lewis, 1978). Besides, the sequences and functions of PcG genes are highly conserved between animals and plants. Several publications recently showed that BPCs can interact with proteins belonging to PRC1 and PRC2, suggesting that it could be a mechanism to repress the expression of their target genes (Wanke et al., 2011; Mu et al., 2017; Xiao et al., 2017). Our results provide further insights into the connection between BPCs and PRC members for the regulation of target genes. We suggest that BPCs of class II and SVP recruit LHP1 and act redundantly with the class I members to establish and maintain H3K27me3 repressive mark on the regulatory region of *STK* in reproductive tissue. In fact, we registered a reduction of

H3K27me3 in the *bpcV* mutant. In the *lhp1* background, we detected increased levels of *STK*, moreover its expression is localized also in the inflorescences and in the floral meristems as well as in the first floral buds. In contrast to our results in the *bpcV*, no signal was detected in other floral structures at maturity, thus suggesting that BPCs of class I and II might repress *STK* expression during flower development also via other mechanisms that do not involve LHP1 activity.

Recently, several BPC targets have been discovered. Most of them are also associated with PRC mediated silencing: the KNOX gene *BREVIPEDICELLUS* (*BP*) is repressed by BPCs throughout flower development (Simonini et al., 2012). The expression of *BP* is directly regulated by the recruitment of the EMBRYONIC FLOWER (EMF) complex by ASYMMETRIC LEAVES 1 and 2 (*AS1* and *AS2*), which triggers H3K27me3 deposition (Lodha et al., 2013). *BP* was also identified in our computational analysis of regions enriched in binding sites for MADS-domain and BPC family members and resulted decorated with H3K27me3 marks (Data S3). Also, *FUS3* has recently been characterised as a BPC target (Roscoe et al., 2019; Wu et al., 2019) and already reported to be a target of PRCs (Makarevich et al., 2006; Zhang et al., 2007a; Bouyer et al., 2011; Yang et al., 2013; Xiao et al., 2017). Interestingly, our computational analysis showed that its regulatory region could also be bound by MADS-domain factors, suggesting a possible conserved mechanism for target regulation.

The MADS-domain factor SVP interacts with LHP1 and is required to recruit the PRC1 factor to the promoter of *SEPALLATA3* (*SEP3*), acting as a pioneer factor (Liu et al., 2009). In accordance to this hypothesis, H3K27me3 deposition on the *SEP3* locus is reduced in *lhp1* background.

Our ChIP assays confirm binding of LHP1 to the 3' end of *STK*, indicating a direct regulation of the homeotic gene by this factor. Interestingly, we previously reported that SVP binds the 3'UTR of many of its targets, among which *STK* (Gregis et al., 2013). Therefore, it is likely that SVP recruits LHP1 on the *STK* locus and subsequently repress the expression of the ovule identity gene via PRC2 recruitment, as previously shown for *SEP3* (Liu et al., 2009).

PRC2 components are required for H3K27me3 deposition to the target locus (Wang et al., 2016). Three different PRC2 complexes regulate plant development by targeting a subset of genes. LHP1 has been reported to associate with several PRC2 members (Derkacheva et al., 2013; Wang et al., 2016), to mediate their recruitment to the target locus. Previously was shown that the EMF

complex played a role in the repression of *AGAMOUS* (*AG*) and *SEP3* in the flower (Yoshida et al., 2001; Kinoshita et al., 2001; Chanvivattana et al., 2004; Calonje et al., 2008). Notably, Derkacheva et al. (2013) reported that LHP1 is directly associated with the EMF complex. We previously characterised *SEP3* and *AG*, as targets of BPCs and SVP (Gregis et al., 2009; Simonini et al., 2012); as matter of fact, the upregulation of *AG* registered in the *lhp1* single mutant is increased in the *lhp1 bpc4 bpc6* triple mutant, confirming a LHP1-class II BPCs interplay in seedlings (Hecker et al., 2015).

Considering all these observations it is tempting to speculate that BPCs and SVP might regulate *STK* expression by the recruitment of LHP1. Then LHP1 as PRC1 member could interact with the EMF complex to mediate the correct deposition of the H3K27me3. Furthermore, LHP1 could assure the maintenance and the spreading of the repressor marks on the *STK* locus throughout flower development (Figure 9).

A general regulatory mechanism in plants

Understanding the molecular mechanisms through which BPCs and SVP containing complexes act, is important, since it is likely that the mechanism by which these factors regulate *STK* can be extended to many other genes during plant development. This is based on the following observations: (i) many genes contain both C-boxes and CARG-boxes in their putative promoter regions; (ii) BPCs are ubiquitously expressed in plants while MADS-domain factors are specifically expressed in all the fundamental developmental stages; and (iii) combination of *bpc* alleles showed pleiotropic phenotypes (Monfared et al., 2011). Furthermore, Berger et al. (2011) identified three cis-elements required for *LEAFY COTYLEDON2* (*LEC2*) repression: C-boxes, CARG-boxes and PRE-like elements, corroborating the idea that the understanding of the synergistic interaction between MADS-domain factors and BPCs is an important key to decode gene regulation in plants. Several key developmental factors that are worth to be tested as putative direct targets of both MADS-domain and BPC factors are reported in Data S3. In fact, they were identified in our computational analysis of regions enriched in binding sites for both MADS and BPC family members. Notably, the analyses of ChIP-seq data available for selected MADS-domain and BPCs in vegetative tissues reveal an highly significant levels of overlap of their binding

profiles in vivo, which suggest a cooperative role of BPCs-MADS factors also during vegetative stages that will be interesting to investigate.

The regulatory mechanism through which BPCs act is of course not restricted to *Arabidopsis*. Several GAGA binding proteins have been discovered in crops and several targets have already been characterised (Sangwan and O'Brian, 2002; Santi et al., 2003; Meister et al., 2004; Gong et al., 2018). Therefore, a better understanding of the mechanisms by which these factors act in *Arabidopsis* may provide knowledge to be used for future crop improvement.

Experimental procedures

Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia was used in this study; the plants were directly sown on soil and kept under short-day conditions for 2 weeks (22°C, 8 h light and 16 h dark) and then moved to long-day conditions (22°C, 16 h light and 8 h dark). The *agl24 svp ap1-12* triple mutant and the *pSVP:SVP-GFP svp* line were previously described by Gregis et al. (2008; 2009); genotyping of the *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutants was done according to Simonini and Kater (2014) and Monfared et al. (2011). Seeds from the *lhp1* (previously named *tfl2-1* (Larsson et al., 1998)), *arf2-8* and *stk* mutant in Columbia background were obtained from the Nottingham Arabidopsis Stock Centre.

Generation of quintuple mutants and marker lines

The *bpc 1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant was obtained by crossing the *bpc 1-2 bpc2 bpc3* triple mutant (Simonini and Kater, 2014) and *bpc4 bpc6* double mutant (Monfared et al., 2011); the *pSVP:SVP-GFP svp bpc1-2 bpc2 bpc3 bpc4 bpc6* was obtained crossing the line previously described by Gregis et al. (2009) and the *bpcV*.

Generation of 35S:STK line

Arabidopsis plants were transformed with the chimeric gene construct in which the CDS of *STK* was fused to the Cauliflower mosaic virus (CaMV) 35S promoter (Favaro et al., 2003) using the *Agrobacterium tumefaciens*-mediated floral dip method (Clough and Bent, 1998). Transformant

plants were sown on MS medium and selected by hygromycin (20 mg/L) resistance; presence of the construct was assessed by genotyping and analysis of *STK* expression.

***STK* promoter constructs and plant transformation**

The mutated version of the *STK* promoter (*pSTK_CARGm*) was synthesised by Twin Helix. The synthetic DNA fragment, like the wild-type version of the *STK* promoter, were cloned in pUC57-Simple (GenScript). The two fragments were digested with *AccI* and *KpnI* and cloned in pDONR207 entry clone (Invitrogen), and successively into pGWB3 binary vector containing the GUS reporter gene. *Arabidopsis* plants were transformed with these constructs using the *Agrobacterium tumefaciens*-mediated floral dip method (Clough and Bent, 1998). Transformant plants were sown on MS plates and selected by hygromycin (20 mg/L) resistance; presence of the construct was assessed by PCR.

GUS staining

GUS assays were performed as described previously by Liljegren et al. (2000). The samples were mounted in lactic acid and subsequently observed using a Zeiss Axiophot D1 microscope equipped with differential interference contrast optics. Images were captured on an Axiocam MRc5 camera (Zeiss) using the Axiovision program (version 4.1).

***In-situ* hybridisation assay**

Arabidopsis flowers were collected, fixed and embedded in paraffin as described by Huijser et al. (1992). Plant tissue sections were probed with *STK* antisense RNA, described in Brambilla et al. (2007); *STK*-sense and H4 histone gene were used as controls (Fobert et al., 1994). Hybridisation and immunological detection were executed as described previously by Coen et al. (1990).

ChIP assay

ChIP assays were performed as described by Gregis et al. (2009) using for SVP-GFP the commercial antibody GFP:Living Colors full-length (Clontech), and for BPCs of class I a polyclonal

antibody as described by Simonini et al. (2012); HA antibody Anti-HA (Roche) were used as negative control in one of the experiments. Quantitative Real-Time PCR assays were performed to determine the enrichment of the fragments. The detection was performed in triplicate using the iQ SYBR Green Supermix (Bio-Rad) and the Bio-Rad iCycler iQ Optical System (software version 3.0a), with the primers listed in Table S2. ChIP-quantitative Real-Time PCR experiments and relative enrichments were calculated as reported by (Matias-Hernandez et al. (2010)). We employed the following formulas to calculate the fold enrichment: $dCT.tg = CT.i - CT.tg$ and $dCT.nc = CT.i - CT.nc$. $CT.tg$ is target gene mean value, $CT.i$ is input DNA mean value, and $CT.nc$ is ACTIN 7 (negative control) mean value: $dCT.tg = CT.i - CT.tg$ and $dCT.nc = CT.i - CT.nc$. The propagated error values of these CTs are calculated using $dSD.tg = \sqrt{((SD.i)^2 + (SD.tg)^2)/n}$ and $dSD.nc = \sqrt{((SD.i)^2 + (SD.nc)^2)/n}$, n = number of replicate per sample. Fold-change over negative control was calculated finding the “delta delta CT” of the target region as follows: $ddCT = dCT.tg - dCT.nc$ and $ddSD = \sqrt{(dSD.tg)^2 + (dSD.nc)^2}$. The transformation to linear “fold-change” values is obtained as follows: $FC = 2^{(ddCT)}$ and $FC.error = \ln(2) * ddSD * FC$. All the experiments were performed in three biological replicates.

ChIP-based analysis of H3K27me3 histone modification

For ChIP-based analysis of histone modifications, the following antibodies were used for immunoprecipitation: Anti-H3K27me3 Rabbit Polyclonal Antibody (Merck 07-449) and Rabbit anti-histone H3 (Sigma-Aldrich H0164). 0,8 mg of grinded and fixed material from unfertilized flowers from wild-type and *bpcV* mutant was collected. ChIP experiments were performed in a modified version of a previously reported protocol (Mizzotti et al., 2014). The quantitative Real-Time PCR assay was conducted in triplicate on four different biological replicates, with three technical replicates for each sample, and was performed in a Bio-Rad iCycler iQ optical system (software version 3.0a). Quantitative Real-Time PCR assays were performed on input and immunoprecipitated samples and % of input was calculated. The signal obtained after precipitation with anti-H3K27me3 antibody (as indicated in Figure 6b) was normalized to actin levels. *AGAMOUS* region was used as a reference as it carries the H3K27me3 mark (Li et al., 2015). Relative enrichment of *AT2G22560* was included as negative control for the H3K27me3 mark (Li et al., 2015). Sequences of oligonucleotides used for ChIP analyses are listed in Table S2.

Yeast two-hybrid assay

The two-hybrid assays were performed at 28°C in the yeast strain AH109 (Clontech). The coding sequences of *BPC4*, *BPC6* and *SVP* were cloned into pDONR207 (Life Technologies) and successively transferred to the Gateway vector GAL4 system (pGADT7 and pGBKT7; Clontech). Yeast two-hybrid assays were performed on selective yeast synthetic dropout medium lacking Leu, Trp, Ade, and His supplemented with different concentrations of 3-aminotriazole (1, 2.5, and 5 mM of 3-AT).

BiFC assay

The *BPC4*, *BPC6* and *SVP* coding sequences were first cloned into pDONR207 (Life Technologies) and subsequently transferred to the pYFPN43 and pYFPC43 vectors by Gateway recombination; while the *AP1* coding sequence was cloned into pDONR207 (Invitrogen) and then transferred to pB7RWG2, purchased from the Flanders Interuniversity Institute for Biotechnology (Gent, Belgium); the previously described formation of VERDANDI-VALKYRIE heterodimers was used as positive control, whereas VERDANDI-VERDANDI combination was used as negative control (Figure S4A; Mendes et al. (2016); all the controls are reported in Figures S3, S4 and S5. BiFC assays were performed injecting *Agrobacterium* expressing viral suppressor p19/experimental constructs as described by Belda-Palazón et al. (2012). The abaxial surfaces of infiltrated tobacco (*Nicotiana benthamiana*) leaves were imaged 3 days after inoculation.

Co-Immunoprecipitation (Co-IP) Protocol

The coding sequences of *BPC4*, *BPC6* and *SVP* were cloned into pDONR221 and then transferred to pB7RWG2 and pB7FWG2, both purchased from the Flanders Interuniversity Institute for Biotechnology (Gent, Belgium). *Nicotiana benthamiana* leaves were infiltrated with *Agrobacterium tumefaciens*, as previously described. 4 days after infiltration, leaf disks (16 mm diameter) were collected and homogenised in 1 ml of immunoprecipitation (IP) buffer (30 mM HEPES-KOH pH 8.0, 200 mM NaCl, 60 mM KOAc, 10 mM MgOAc, 0.5% [v/v] Nonidet P-40 and proteinase inhibitor cocktail [cOmplete™, COEDTAF-RO, Roche]). Samples were incubated in ice for 15 min to allow membrane solubilisation and subjected to a centrifugation step (10 min at

16,000 g). Supernatants were incubated (2 h, at 4°C) with 20 µl RFP-Trap®_MA (ChromoTek) or GFP-Trap®_MA (ChromoTek). Beads were then washed 3 times for 10 min with 1 ml of IP buffer and eluted with Laemmli sample buffer. Protein samples were fractionated on SDS–PAGE (10% [w/v] acrylamide (Schägger and von Jagow, 1987)) and then transferred to polyvinylidene difluoride (PVDF) membranes. Filters were immuno-decorated with specific antibodies; the Coomassie Brilliant Blue (CBB) staining of the gel was performed as loading control. The anti-GFP antibody was purchased from Thermo Fisher Scientific while the anti-RFP antibody was obtained from ChromoTek.

Gene expression analysis

Quantitative Real-Time PCR experiments were performed using cDNA obtained from inflorescences. Total RNA was extracted using lithium chloride. The Ambion TURBO DNA-free DNase kit was used to remove genomic DNA contaminations, according to the manufacturer's instructions (<http://www.ambion.com/>). The ImProm-IITM reverse transcription system (Promega) was used to retrotranscribe the treated RNA. Transcripts were detected using a Sybr Green Assay (iQ SYBR Green Supermix; Bio-Rad) using *UBIQUITIN* as a reference gene. Assays were done in triplicate using a Bio-Rad iCycler iQ Optical System (software version 3.0a). The enrichments were calculated normalising the amount of mRNA against housekeeping gene fragments. The expression of different genes was analysed using specific oligonucleotides primers (Table S2).

Microscopy and imaging

Images of plants, cauline and rosette leaves were acquired using a Canon EOS 6D camera whereas images of siliques were taken using a Leica® MZ 6 stereomicroscope. For in-situ experiments sections were analysed using a Zeiss Axiophot D1 microscope supplied with differential interface contrast (DIC) optics and Axiocam MRc5 camera (Zeiss) using the AXIOVISION program (version 4.4).

Scanning

Seed area size were analysed by using SMART-GRAIN software. ANOVA and post-hoc Tukey HSD (honestly significant difference) test were used for wild-type versus other genotypes comparison.

Computational analyses

DAP-seq peaks data were obtained in the form of narrowpeaks files from http://neomorph.salk.edu/dap_web/pages/index.php. Narrow-peaks files were concatenated and overlapped genomic regions were merged by the means of the bedtools merge utility. Finally, candidate binding regions showing a positive hit for the majority (that is $n/2+1$, if profiles for n family members were available) of the members of a family were retained to form the "consensus" family profile. ChIP-seq peaks for BPC1 (GSE84483), BPC6, SVP (GSE54881), SOC1 (GSE45846), FLC (GSE54881) and FLM (GSE48082) were retrieved directly from their respective entries in the GEO database. Data of selected MADS-box ChIP-seq were chosen based on the tissue in which the experiments were performed: seedling or vegetative tissues as for ChIP-seqs available for both BPC1 and BPC6. Intersection of peaks coordinates were performed using the bedtools intersect program (Quinlan and Hall, 2010) using default parameters; peaks with an overlap of 1 bp were considered coincident. The "-u" option was used in order to collapse peaks showing multiple overlaps. Statistical significance of overlaps was assessed by using the hypergeometric distribution. Annotation of selected DAP-seq peaks was performed by the means of the annotatePeaks program from the Homer suite (Heinz et al., 2010) using the reference TAIR10 annotation. Identification of enriched sequence motifs and identification of closely related motifs from publicly available dataset of were performed by the means of the findMotifsGenome utility in Homer. Functional enrichment analyses were performed by using the web interface of the DAVID suite (Huang et al., 2009).

Accession numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: *STK* (AT4G09960), *BPC1* (AT2G01930), *BPC2* (AT1G14685), *BPC3* (AT1G68120), *BPC4* (AT2G21240), *BPC6* (AT5G42520), *AGL24* (AT4G24540), *SVP* (AT2G22540), *AP1* (AT1G69120), *LHP1*

(AT5G17690), *NETWORK*

2D (AT2G22560), *AGAMOUS* (AT4G18960), *VERDANDI* (AT5G18000), *VALKYRIE* (AT2G24690), *ACTIN7* (AT5G09810) and *UBIQUITIN* (AT4G36800).

Conflict of interest

The authors declare that they have no conflicts of interest.

Funding

This work was supported by the Ministero dell'Istruzione, dell'Università e della Ricerca MIUR, SIR2014 MADSMEC, Proposal number RBSI14BTZR. The PhD fellowships of R.P. and F.C. were supported by the Doctorate School in Molecular and Cellular Biology, Università degli Studi di Milano. R.P. and F.C. were supported by H2020-MSCA-RISE-2015 ExpoSEED Proposal Number: 691109.

Author contributions

V.G. designed the research strategy and the experiments. V.G. and R.P. performed most of the experiments. F.C. performed yeast two-hybrid analyses. T.L. designed and performed CoIP experiments. V.V. did the expression analysis on *lhp1* by quantitative Real-Time PCR and performed bioinformatics analyses, M.C. supervised and performed bioinformatics analyses. I.R.V. obtained the *bpcV* multiple mutant. I.E. performed ChIP experiments on chromatin modifications and analysis on seed size area. V.G., R.P. and M.M.K. interpreted data and wrote the manuscript. All authors discussed the results and commented on the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank V. Borrelli, A. Ravasio, A. Piva (Università degli Studi di Milano) and S. de Folter (Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico) for technical support. S. Masiero, P. Pesaresi, L. Colombo, L. Conti, L. Gianfranceschi (Università degli Studi di

Milano) and R. Battaglia (CREA-GB - Fiorenzuola d'Arda PC) for helpful suggestions and valuable discussions. Part of this work was carried out at NOLIMITS, an advanced imaging facility established by the Università degli Studi di Milano.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Figure S1. *In situ* hybridisation controls.

Figure S2. Selection of 35S:STK T1 lines and morphological analyses on 35S:STK and *bpcV* background.

Figure S3. Yeast two-hybrid assay between BPCs of class II, Co-immunoprecipitation assays using GFP-trap and Bi-molecular fluorescence complementation (BiFC) assay between BPCs of class II, between BPCs of class II and AP1 and between BPC2 and SVP.

Figure S4. Bi-molecular fluorescence complementation (BiFC) assay positive controls and cellular localization of BPC4-RFP, BPC6-RFP, AP1-RFP and SVP-GFP.

Figure S5. Bi-molecular fluorescence complementation (BiFC) assay, negative controls.

Figure S6. Proportion of significantly overlapped DAP-Seq Peaks by TF family.

Figure S7. Venn diagrams of common ChIP-seq peaks between BPC1, BPC6 and a selection of MADS-domain TFs

Table S1. MADS-domain consensus regions identified in *STK* promoter and the designed mutated versions.

Table S2. Sequences of oligonucleotides used in this manuscript.

Data S1. p-values for the intersection of DAP-seq peaks of TFs with DAP-seq peaks of the BPCs family of transcription factors.

Data S2. p-values for the intersection of ChIP-seq peaks of a selection of MADS-domain TFs with ChIP-seq peaks of BPC1 and BPC6.

Data S3. Lists of identified genome-wide binding locations for all the MADS-box and BPCs factors, coincident DAP-seq peaks and functional enrichment analyses of associated target genes.

References

- Balanza, V., Roig-Villanova, I., Di Marzo, M., Masiero, S., and Colombo, L. (2016). Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. *Development* **143**:3372–3381.
- Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., and Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**:1659–1672.

Belda-Palazón, B., Ruiz, L., Martí, E., Tárraga, S., Tiburcio, A. F., Culiáñez, F., Farràs, R., Carrasco, P., and Ferrando, A. (2012). Aminopropyltransferases Involved in Polyamine Biosynthesis Localize Preferentially in the Nucleus of Plant Cells. *PLoS One* **7**:e46907.

Berger, N., and Dubreucq, B. (2012). Evolution goes GAGA: GAGA binding proteins across kingdoms. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**:863–868.

Berger, N., Dubreucq, B., Roudier, F., Dubos, C., and Lepiniec, L. (2011). Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell* **23**:4065–78.

Bouyer, D., Roudier, F., Heese, M., Andersen, E. D., Gey, D., Nowack, M. K., Goodrich, J., Renou, J.-P., Grini, P. E., Colot, V., et al. (2011). Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. *PLoS Genet.* **7**:e1002014.

Brambilla, V., Battaglia, R., Colombo, M., Masiero, S., Bencivenga, S., Kater, M. M., and Colombo, L. (2007). Genetic and Molecular Interactions between BELL1 and MADS Box Factors Support Ovule Development in Arabidopsis. *Plant Cell Online* Advance Access published 2007, doi:10.1105/tpc.107.051797.

Calonje, M., Sanchez, R., Chen, L., and Sung, Z. R. (2008). EMBRYONIC FLOWER1 Participates in Polycomb Group–Mediated AG Gene Silencing in Arabidopsis. *Plant Cell* **20**:277–291.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.-H., Sung, Z. R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**:5263–76.

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**:735–43.

Coen, E. S., Romero, J. M., Doyle, S., Elliott, R., Murphy, G., and Carpenter, R. (1990). floricaula: a homeotic gene required for flower development in antirrhinum majus. *Cell* **63**:1311–22.

de Folter, S., Immink, R. G. H., Kieffer, M., Parenicová, L., Henz, S. R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M. M., et al. (2005). Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. *PLANT CELL ONLINE* **17**:1424–1433.

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgová, I., Mahrez, W., Nanni, P., Bischof, S., Grisse, W., and Hennig, L. (2013a). Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *EMBO J.* **32**:2073–

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgová, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W., and Hennig, L. (2013b). Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *EMBO J.* **32**:2073–85.

Doughty, J., Aljabri, M., and Scott, R. J. (2014). Flavonoids and the regulation of seed size in *Arabidopsis*. *Biochem. Soc. Trans.* **42**:364–369.

Ezquer, I., Mizzotti, C., Nguema-Ona, E., Gotté, M., Beauzamy, L., Viana, V. E., Dubrulle, N., Costa de Oliveira, A., Caporali, E., Koroney, A.-S., et al. (2016). The Developmental Regulator SEEDSTICK Controls Structural and Mechanical Properties of the Arabidopsis Seed Coat. *Plant Cell* **28**:2478–2492.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* **371**:806–808.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., and Borghi, L. (2003). MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. *Plant Cell* **15**.

Fobert, P. R., Coen, E. S., Murphy, G. J., and Doonan, J. H. (1994). Patterns of cell division revealed by transcriptional regulation of genes during the cell cycle in plants. *EMBO J.* **13**:616–24.

Gong, R., Cao, H., Zhang, J., Xie, K., Wang, D., and Yu, S. (2018). Divergent functions of the GAGA-binding transcription factor family in rice. *Plant J.* **94**:32–47.

Gregis, V., Sessa, A., Colombo, L., and Kater, M. M. (2006). AGL24, SHORT VEGETATIVE PHASE and APETALA1 Redundantly Control AGAMOUS during Early Stages of Flower Development in Arabidopsis. *The Plant Cell*. **18**:1373–1382.

Gregis, V., Sessa, A., Colombo, L., and Kater, M. M. (2008). AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in Arabidopsis. *Plant J.* **56**:891–902.

Gregis, V., Sessa, A., Dorca-Fornell, C., and Kater, M. M. (2009). The Arabidopsis floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *Plant J.* **60**:626–637.

Gregis, V., Andrés, F., Sessa, A., Guerra, R. F., Simonini, S., Mateos, J. L., Torti, S., Zambelli, F., Prazzoli, G. M., Bjerkan, K. N., et al. (2013). Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in Arabidopsis. *Genome Biol.*

Hecker, A., Brand, L. H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D.

(2015a). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant Physiol.* Advance Access published 2015, doi:10.1104/pp.15.00409.

Hecker, A., Brand, L. H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D.

(2015b). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant Physiol.* **168**:1013–1024.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and

Glass, C. K. (2010a). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**:576–589.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and

Glass, C. K. (2010b). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**:576–589.

Herrera-Ubaldo, H., Lozano-Sotomayor, P., Ezquer, I., Di Marzo, M., Chávez Montes, R. A., Gómez-

Felipe, A., Pablo-Villa, J., Diaz-Ramirez, D., Ballester, P., Ferrándiz, C., et al. (2019). New roles of NO TRANSMITTING TRACT and SEEDSTICK during medial domain development in Arabidopsis fruits. *Development* **146**:dev172395.

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). Structure of a polycomb response element and in

vitro binding of polycomb group complexes containing GAGA factor. *Mol. Cell. Biol.* **20**:3187–97.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large

gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**:44–57.

Immink, R. G. H., Gadella, T. W. J., Ferrario, S., Busscher, M., and Angenent, G. C. (2002). Analysis of

MADS box protein-protein interactions in living plant cells. *Proc. Natl. Acad. Sci.* **99**:2416–2421.

Immink RG, Posé D, Ferrario S, Ott F Kaufmann K, Valentim FL, de Folter S, van der Wal F, van Dijk AD,

Schmid M, Angenent GC. (2012) Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. *Plant Physiol* **160**(1):433-49.

Mateos JL, Madrigal P, Tsuda K, Rawat V Richter R, Romera-Branchat M, Fornara F, Schneeberger K,

- Krajewski P, Coupland G.** (2015). Combinatorial activities of SHORT VEGETATIVE PHASE and FLOWERING LOCUS C define distinct modes of flowering regulation in Arabidopsis. *Genome Biol.* **16**:31.
- Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J. A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S. R., Tan, G., Baranasic D, Arenillas DJ, Sandelin A, Vandepoele K, Lenhard B, Ballester B, Wasserman WW, Parcy F, Mathelier A.** (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46**:D260–D266.
- Kinoshita, T., Harada, J. J., Goldberg, R. B., and Fischer, R. L.** (2001). Polycomb repression of flowering during early plant development. *Proc. Natl. Acad. Sci. U. S. A.* **98**:14156–61.
- Kooiker, M., CA, A., A, L., PS, M., L, F., MM, K., and L., C.** (2005). BASIC PENTACYSTEINE1, a GA Binding Protein That Induces Conformational Changes in the Regulatory Region of the Homeotic Arabidopsis Gene SEEDSTICK. *PLANT CELL ONLINE* **17**:722–729.
- Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K.** (2003). Arabidopsis TERMINAL FLOWER 2 Gene Encodes a Heterochromatin Protein 1 Homolog and Represses both FLOWERING LOCUS T to Regulate Flowering Time and Several Floral Homeotic Genes. *Plant Cell Physiol.* **44**:555–564.
- Lafos, M., Kroll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., and Schubert, D.** (2011). Dynamic Regulation of H3K27 Trimethylation during Arabidopsis Differentiation. *PLoS Genet.* **7**:e1002040.
- Larsson, A. S., Landberg, K., and Meeks-Wagner, D. R.** (1998). The TERMINAL FLOWER2 (TFL2) gene controls the reproductive transition and meristem identity in Arabidopsis thaliana. *Genetics* **149**:597–605.
- Lewis, E. B.** (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**:565–570.
- Li, C., Chen, C., Gao, L., Yang, S., Nguyen, V., Shi, X., Siminovitch, K., Kohalmi, S. E., Huang, S., Wu, K., et al.** (2015). The Arabidopsis SWI2/SNF2 Chromatin Remodeler BRAHMA Regulates Polycomb Function during Vegetative Development and Directly Activates the Flowering Repressor Gene SVP. *PLOS Genet.* **11**:e1004944.
- Liljegren, S. J., Ditta, G. S., Eshed, Y., Savidge, B., Bowman, J. L., and Yanofsky, M. F.** (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature* **404**:766–770.
- Liu, C., Xi, W., Shen, L., Tan, C., and Yu, H.** (2009). Regulation of Floral Patterning by Flowering Time

Genes. *Dev. Cell* **16**:711–722.

Lodha, M., Marco, C. F., and Timmermans, M. C. P. (2013). The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes Dev.* **27**:596–601.

Losa, A., Colombo, M., Brambilla, V., and Colombo, L. (2010). Genetic interaction between AINTEGUMENTA (ANT) and the ovule identity genes SEEDSTICK (STK), SHATTERPROOF1 (SHP1) and SHATTERPROOF2 (SHP2). *Sex. Plant Reprod.* **23**:115–121.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Köhler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep.* **7**:947–952.

Malley, R. C. O., Carol, S., Song, L., Galli, M., Ecker, J. R., Malley, R. C. O., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., et al. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape Resource Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.

Martinez, G. J., and Rao, A. (2012). Cooperative Transcription Factor Complexes in Control. *Science* (80-.). **338**:891–892.

Matias-Hernandez, L., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., Kater, M. M., and Colombo, L. (2010). VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis. *Plant Cell* **22**:1702–1715.

Meister, R. J., Williams, L. A., Monfared, M. M., Gallagher, T. L., Kraft, E. A., Nelson, C. G., and Gasser, C. S. (2004). Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER NO OUTER promoter. *Plant J.* **37**:426–438.

Mendes, M. A., Guerra, R. F., Castelnovo, B., Silva-Velazquez, Y., Morandini, P., Manrique, S., Baumann, N., Groß-Hardt, R., Dickinson, H., and Colombo, L. (2016). Live and let die: a REM complex promotes fertilization through synergid cell death in *Arabidopsis*. *Development* **143**:2780–2790.

Mizzotti, C., Ezquer, I., Paolo, D., Rueda-Romero, P., Guerra, R. F., Battaglia, R., Rogachev, I., Aharoni, A., Kater, M. M., Caporali, E., et al. (2014). SEEDSTICK is a Master Regulator of Development and Metabolism in the Arabidopsis Seed Coat. *PLoS Genet.* **10**:e1004856.

Monfared, M. M., Simon, M. K., Meister, R. J., Roig-Villanova, I., Kooiker, M., Colombo, L., Fletcher, J. C.,

- and Gasser, C. S. (2011). Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *Plant J.* **66**:1020–1031.
- Mu, Y., Zou, M., Sun, X., He, B., Xu, X., Liu, Y., Zhang, L., and Chi, W. (2017a). BASIC PENTACYSTEINE proteins repress Absciscic Acid INSENSITIVE 4 expression via direct recruitment of the polycomb-repressive complex 2 in arabidopsis root development. *Plant Cell Physiol.* **58**:607–621.
- Mu, Y., Zou, M., Sun, X., He, B., Xu, X., Liu, Y., Zhang, L., and Chi, W. (2017b). Basic Pentacysteine Proteins Repress *Absciscic Acid Insensitive4* Expression via Direct Recruitment of the Polycomb-Repressive Complex 2 in *Arabidopsis* Root Development. *Plant Cell Physiol.* **58**:pcx006.
- O'Malley, R. C., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A., and Ecker, J. R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.
- Pelaz, S., Gustafson-Brown, C., Kohalmi, S. E., Crosby, W. L., and Yanofsky, M. F. (2002). APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J.* **26**:385–394.
- Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., and Yanofsky, M. F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* **424**:85–88.
- Posé, D., Verhage, L., Ott, F., Yant L., Mathieu, J., Angenent G.C., Immink R.G., Schmid M. (2013). Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature* **21**;503(7476):414-7.
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842.
- Roscoe, T. J., Vaissayre, V., Paszkiewicz, G., Clavijo, F., Kelemen, Z., Michaud, C., Lepiniec, L., Dubreucq, B., Zhou, D.-X., and Devic, M. (2019). Regulation of *FUSCA3* Expression During Seed Development in Arabidopsis. *Plant Cell Physiol.* **60**:476–487.
- Sangwan, I., and O'Brian, M. R. (2002). Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA. *Plant Physiol.* **129**:1788–94.
- Santi, L., Wang, Y., Stile, M. R., Berendzen, K., Wanke, D., Roig, C., Pozzi, C., Müller, K., Müller, J., Rohde, W., et al. (2003). The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*. *Plant J.* **34**:813–826.

- Schägger, H., and von Jagow, G.** (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368–79.
- Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R. J.** (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development* **133**:251–61.
- Shanks, C. M., Hecker, A., Cheng, C.-Y., Brand, L., Collani, S., Schmid, M., Schaller, G. E., Wanke, D., Harter, K., and Kieber, J. J.** (2018). Role of *BASIC PENTACYSTEINE* transcription factors in a subset of cytokinin signaling responses. *Plant J.* **95**:458–473.
- Simonini, S., and Kater, M. M.** (2014). Class I BASIC PENTACYSTEINE factors regulate HOMEBOX genes involved in meristem size maintenance. *J. Exp. Bot.* **65**:1455–1465.
- Simonini, S., Roig-Villanova, I., Gregis, V., Colombo, B., Colombo, L., and Kater, M. M.** (2012). BASIC PENTACYSTEINE Proteins Mediate MADS Domain Complex Binding to the DNA for Tissue-Specific Expression of Target Genes in *Arabidopsis*. *Plant Cell* **24**:4163–4172.
- Theune, M. L., Bloss, U., Brand, L. H., Ladwig, F., and Wanke, D.** (2019). Phylogenetic Analyses and GAGA-Motif Binding Studies of BBR/BPC Proteins Lend to Clues in GAGA-Motif Recognition and a Regulatory Role in Brassinosteroid Signaling. *Front. Plant Sci.* **10**:466.
- Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.-L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R. A., Coupland, G., and Colot, V.** (2007). Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27. *PLoS Genet.* **3**:e86.
- Wang, H., Liu, C., Cheng, J., Liu, J., Zhang, L., and He, C.** (2016). Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements Advance Access published 2016, doi:10.1371/journal.pgen.1005771.
- Wanke, D., Hohenstatt, M. L., Dynowski, M., Bloss, U., Hecker, A., Elgass, K., Hummel, S., Hahn, A., Caesar, K., Schleifenbaum, F., et al.** (2011). Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS One* **6**:e16070.
- Wu, J., Petrella, R., Dowhanik, S., Gregis, V., and Gazzarrini, S.** (2019). Spatiotemporal restriction of FUSCA3 expression by class I BPC promotes ovule development and coordinates embryo and

endosperm growth. *bioRxiv* Advance Access published April 19, 2019, doi:10.1101/612408.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., Song, C., Zhuang, M., Klasfeld, S., He, C., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nat. Genet.* **49**:1546–1552.

Yang, J., Lee, S., Hang, R., Kim, S.-R., Lee, Y.-S., Cao, X., Amasino, R., and An, G. (2013). OsVIL2 functions with PRC2 to induce flowering by repressing *O s LFL 1* in rice. *Plant J.* **73**:566–578.

Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z. R., and Takahashi, S. (2001). EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell* **13**:2471–81.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., and Jacobsen, S. E. (2007a). Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in *Arabidopsis*. *PLoS Biol.* **5**:e129.

Zhang, X., Germann, S., Blus, B. J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S. E. (2007b). The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat. Struct. & Mol. Biol.* **14**:869.

Figures legends.

Figure 1. Mutation of BPCs of class I and class II affects *STK* expression in the flower.

In-situ hybridisation on wild-type [(a) and (b)], *bpc1-2 bpc2 bpc3* [(c) and (d)], *bpc4 bpc6* [(e) and (f)] and *bpcV* [(g) and (h)] inflorescences using a *STK*-specific antisense probe. IM: inflorescence meristem; P: petal; numbers represent flower stages. Scale bars=50 µm.

Figure 2. Overexpression of *STK* affects vegetative and reproductive development.

(a) From left to right: wild-type, *bpcV* and 35S:STK plants; plants were photographed six weeks after sowing; scale bars=1 cm. (b) Fruit morphology and length in wild type, *bpcV* and 35S:STK (from top to bottom); scale bars=1.5 mm. (c) Average seeds area size of wild-type, *arf2-8* (Schruff et al., 2006), *stk* (Pinyopich et al., 2001), 35S:STK and *bpcV*; error bars represent the standard error mean of replicates; ANOVA and post-hoc Tukey HSD (honestly significant difference) test were used, **P < 0.01 for wild-type versus other genotypes comparison. In the lower row, seeds of the analysed genotypes are shown.

Figure 3. Class II BPCs interact with SVP *in vivo*.

(a) Yeast two-hybrid interaction assay for SVP and BPCs of class II: positive interactions on selective media –W-L-H +5mM 3-AT.

(b) Co-immunoprecipitation assays. *Nicotiana benthamiana* leaves were infiltrated with constructs carrying SVP-GFP together with BPC4-RFP and BPC6-RFP, as described in experimental procedures. Immunoprecipitation step was performed using RFP-trap on total protein leaf extract. Samples were probed with GFP and RFP antibodies. S/N: supernatant; IP: immunoprecipitation.

(c) Bi-molecular fluorescence complementation (BiFC) assay. *Nicotiana benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second column yellow fluorescence and the merging in the bright field were shown, respectively.

(d) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions and AP1-RFP construct. In the first, the second and the third column yellow fluorescence, red fluorescence and the merging between the two channels in the bright field were shown, respectively. Scale bars=40 μ m.

Figure 4. ChIP experiments on different mutant backgrounds.

ChIP experiments on different mutant backgrounds. (a) Schematic diagram of the *STK* locus indicating the regions analysed by chromatin immunoprecipitation (ChIP; black bars). Black boxes, exons; white boxes, promoters and introns; asterisks, C-boxes; grey boxes, CArG-boxes; scale bar=500 bp. (b) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *svp ap1-12 agl24*, wild-type (as a positive control), and *bpc1-2 bpc2 bpc3* (as a negative control) testing the C-12, B and NC box regions. Antibodies against BPCs of class I were used. (c) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP svp bpcV*, *pSVP:SVP-GFP svp* (as a positive control) and wild-type (as a negative control), testing C-12, B and NC box regions. For the IP, commercial antibodies against GFP were used. Error bars represent the propagated error value using three replicates. ChIP results of one representative experiment are shown. Positive binding site fragments were considered only if they were enriched compared with the controls in at least three independent experiments.

Figure 5. Mutation of CArG-boxes interferes with SVP and class I BPCs binding to *STK* promoter.

(a) Schematic representation of the *STK* promoter versions generated: dark grey squares represent CArG-boxes wild-type and mutated (crossed). (b)-(g) GUS staining on inflorescences from *pSTK:GUSwt* (b-d) and *pSTK_CArGm:GUS* (e-g): whole inflorescence [(b) and (e)]; mature flower [(c) and (f)]; inflorescence meristem (IM), floral meristems (FM) and young flowers [(d) and (g)]; scale bars in (c), (d), (f) and (g)=100 μ m. (h) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP pSTK_CArGm svp* showing deregulation of the reporter and *pSTK_CArGm* as a negative control, testing wild-type region, mutated region and NC box. For the IP, antibodies against GFP have been used. (i) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP pSTK_CArGm svp* showing correct expression of the reporter and *pSTK_CArGm* as a negative control, testing wild-type region, mutated region and NC box. For the IP, antibodies against GFP have been used. (l) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP pSTK_CArGm* testing wild-type region, mutated region and NC box. For the IP, antibodies against

Class I BPCs have been used; for negative control commercial antibodies against HA was used. Error bars represent the propagated error value using three replicates. ChIP results of one representative experiment are shown. Positive binding site fragments were considered only if they were significantly enriched compared to the controls in at least three independent experiments.

Figure 6. Epigenetic regulation of *STK*.

(a) Schematic representation of the *STK* genomic region tested in ChIP assay. Black boxes indicate exonic regions. Black bars indicate the regions analysed by chromatin immunoprecipitation (ChIP); region1 is located in the H3K27me3 – enriched region published by (Li et al. 2015) spanning -2627 upstream *STK*-transcriptional start site to +2050 pb downstream *STK*-transcriptional start site, whereas region 2 is localized 3 pb downstream the stop codon of the gene. Black arrow indicates the *STK*-transcription start site. Scale bar= 500 bp.

(b) ChIP-quantitative Real-Time PCR determining the levels of H3K27me3 across the *STK* locus in inflorescence tissue. Quantitative Real-Time PCR quantification of *STK* sequences in precipitated chromatin was used to infer the methylation of histone H3 at lysine 27 (H3K27me3) and histone H3 density. Ct values were used to calculate the IP/IN signal. ChIP enrichments are presented as the percentage (%) of bound/input signal normalized to actin levels in the relative regions. We tested the efficiency of IP on histone modifications by quantifying the presence of the H3K27me3 mark in AG region which carries the mark H3k27me3, reported in Li et al. (2015). H3K27me3 mark in AT2G22560 was used as negative control for H3K27me3 mark (Li et al. 2015). The data were normalized to *actin*, with error bars indicating standard deviations based on three independent technical replicates. Four, independent ChIP experiments were performed and similar results were obtained.

Figure 7. LHP1 directly regulates *STK* during flower development.

(a)-(d) *In-situ* hybridisation on wild-type [(a) and (b)] and *lhp1* inflorescences [(c) and (d)] using a *STK*-specific antisense probe (Brambilla et al., 2007). IM: inflorescence meristem; FM: floral meristem; numbers represent flower stages; scale bars=50 μ m.

(e) Expression analysis of *STK* by quantitative Real-Time PCR in *lhp1* and wild-type inflorescences. The expression of *STK* was normalized to that of ubiquitin and the expression level in wild-type was set to 1. Asterisk indicates $P < 0.05$ in a Student's t-test.

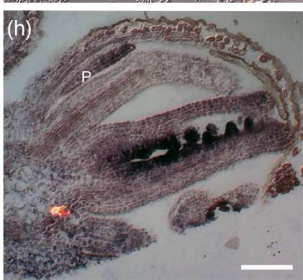
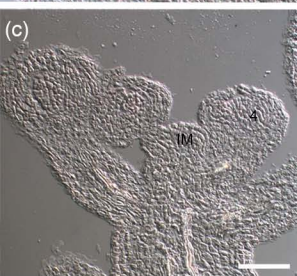
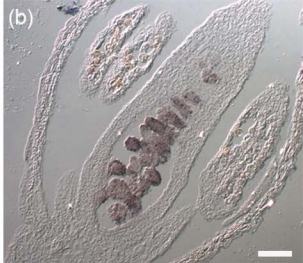
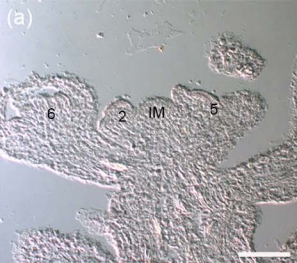
(f) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pLHP1:LHP1-GFP* (Kotake et al, 2003) and wild-type (as a negative control), testing the Region 1 and Region 2 (Figure 6a). For the IP, commercial antibodies against SVP were used. Error bars represent the propagated error value using three replicates. ChIP results of one representative experiment are shown. Positive binding site fragments were considered only if they were significantly enriched compared with the controls in at least three independent experiments.

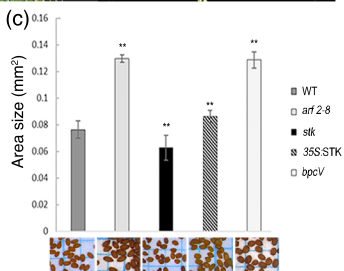
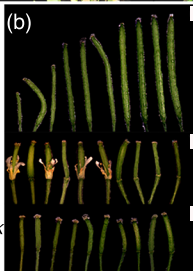
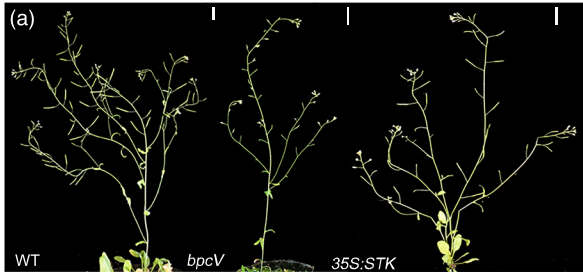
Figure 8. Analysis of BPC and MADS-box transcription factor families binding sites.

Venn diagram displaying the number of DAP-seq peaks and the common number of peaks associated to the BPC and MADS transcription factor families according to our analysis of the data by O'Malley et al (see Experimental procedures). Enriched motifs, as recovered by Homer (p -value $\leq 1e-30$), are displayed underneath.

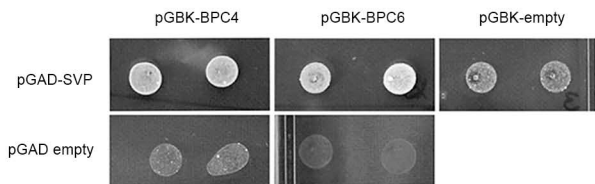
Figure 9. Model of the protein complex formed to represses gene expression during flower development.

BPCs and SVP bind C-boxes (in dark purple) and CARG-boxes (in light purple) respectively, and recruit LHP1 to a subset of gene loci.

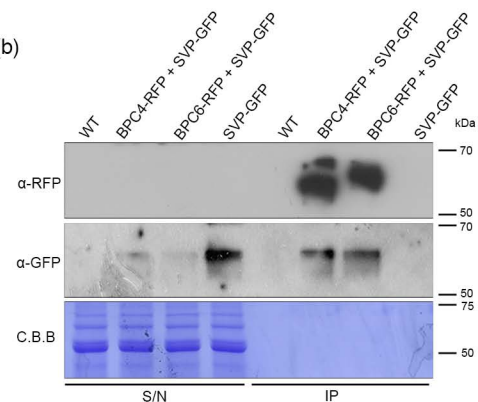




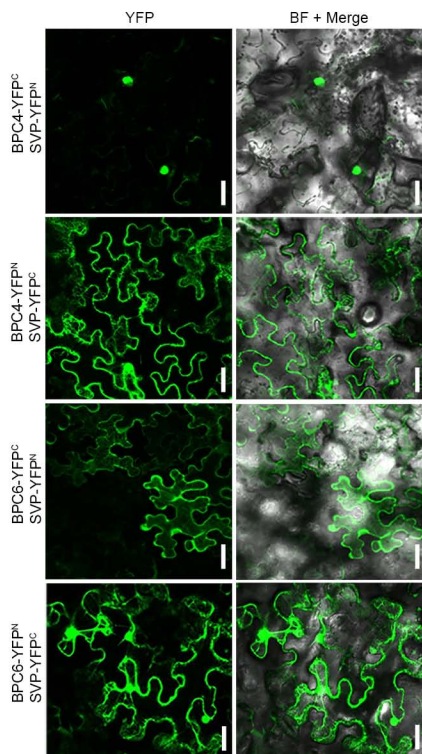
(a)



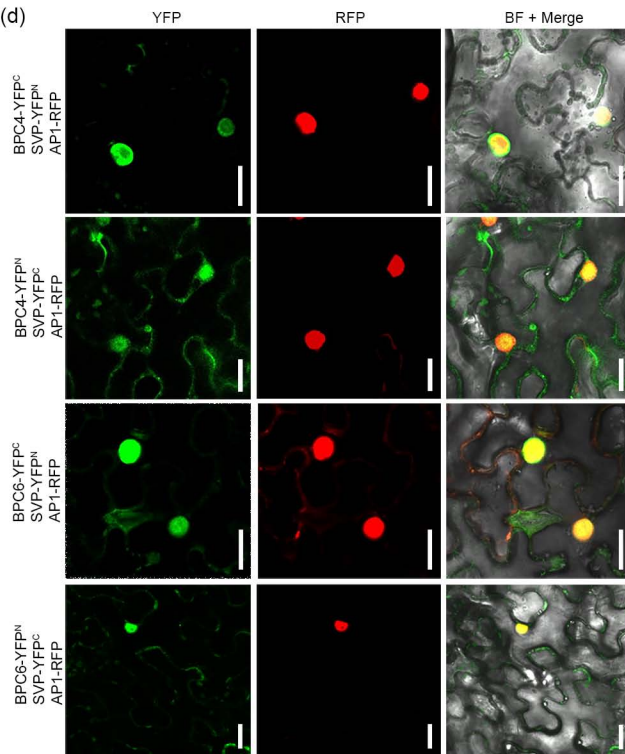
(b)



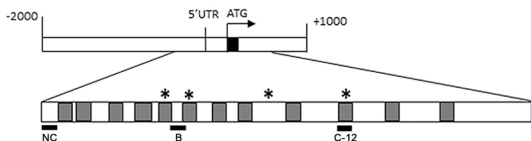
(c)



(d)

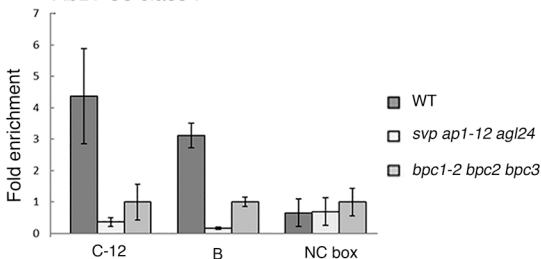


(a)



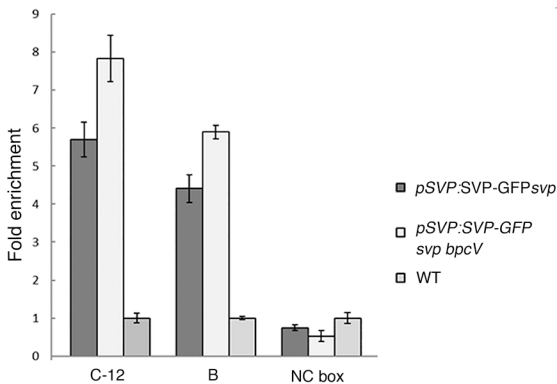
(b)

AbBPCs class I

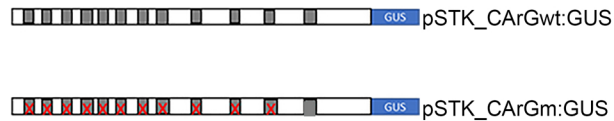


(c)

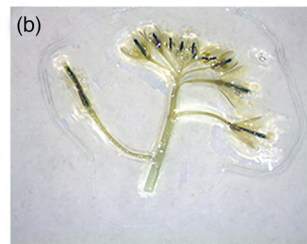
AbGFP



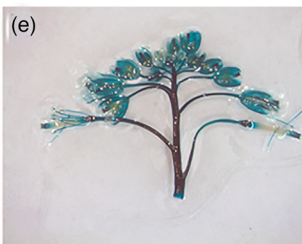
(a)



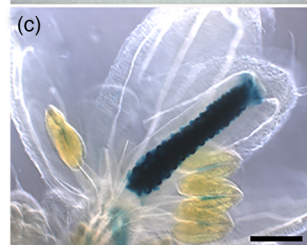
(b)



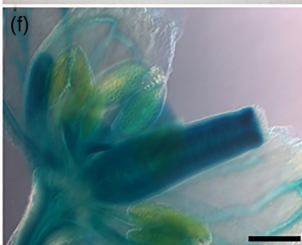
(e)



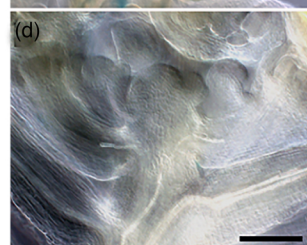
(c)



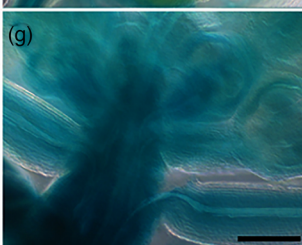
(f)



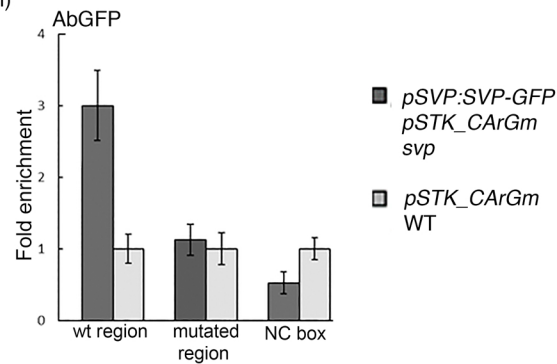
(d)



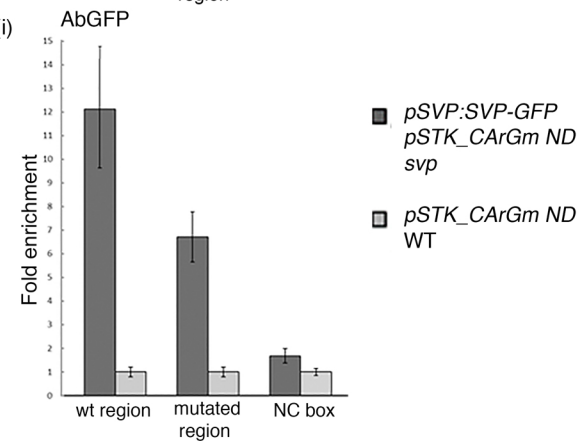
(g)



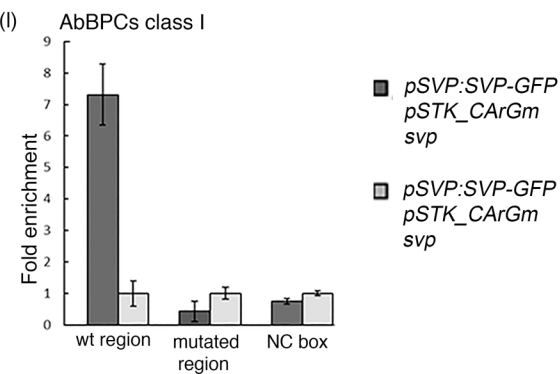
(h)



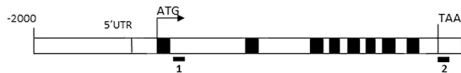
(i)



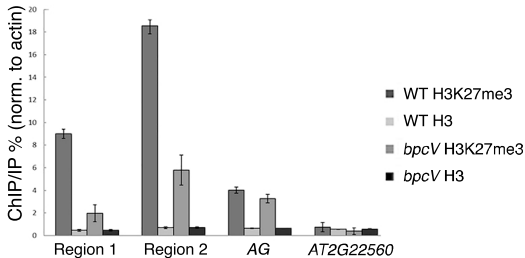
(l)

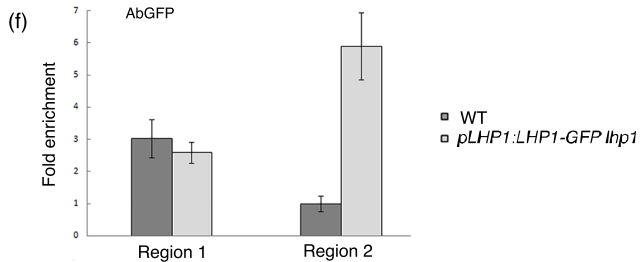
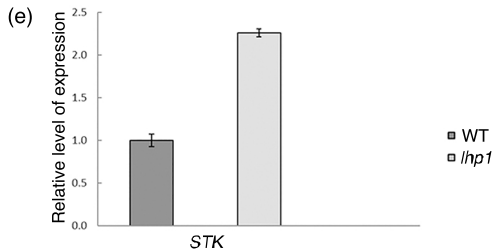
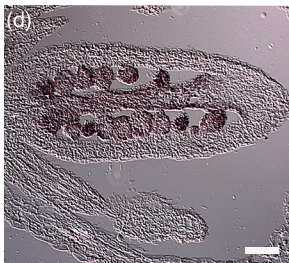
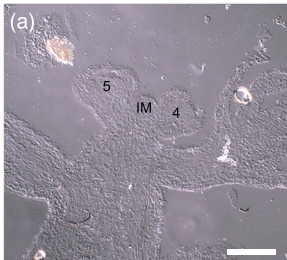


(a)



(b)





BPC

MADS



	Motif	P-value	% Targets
1	TTCTCTCTCT	1e-5688	74.45
2	CTTCCTTC	1e-385	29.96
3	GGTGGTGAGAGG	1e-286	10.59
4	GTGTGTGTGTGT	1e-107	4.59
5	TATATATA	1e-36	15.85
6	ATAAATAAATAA	1e-33	9.22
7	AGGGAAGA	1e-31	4.79

	Motif	P-value	% Targets
1	TCTCTTTC	1e-201	83.90
2	CCAAATSGGAA	1e-167	56.06
3	CCCSTTTT	1e-36	35.23

	Motif	P-value	% Targets
1	CCAAAGSAGG	1e-3221	64.10
2	AAAGGSA	1e-831	71.69
3	TCGGGTAA	1e-281	32.54
4	CCCGGAAAGGAA	1e-129	2.18
5	AAAATGGSCA	1e-91	16.91
6	CCAAAGGTG	1e-61	5.93
7	AATGTGATC	1e-40	25.09

●●● H3K27me3

